AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD025361

This article is a companion to *Jeong et al.* [2016] doi:10.1002/2016JD025404.

Key Points:
- Meteorological observations used to select site-specific WRF model schemes
- CO measurements compared with WRF-STILT model predictions
- Model biases and random errors estimated for measurement sites

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Jeong,
sjeong@lbl.gov

Citation:
Bagley, J. E., et al. (2017), Assessment of an atmospheric transport model for annual inverse estimates of California greenhouse gas emissions, *J. Geophys. Res. Atmos.*, *122*, 1901–1918, doi:10.1002/2016JD025361.

Received 13 MAY 2016
Accepted 16 JAN 2017
Accepted article online 19 JAN 2017
Published online 2 FEB 2017

Assessment of an atmospheric transport model for annual inverse estimates of California greenhouse gas emissions

Justin E. Bagley[1], Seongeun Jeong[1], Xinguang Cui[1], Sally Newman[2], Jingsong Zhang[3], Chad Priest[3], Mixtli Campos-Pineda[3], Arlyn E. Andrews[4], Laura Bianco[4], Matthew Lloyd[5], Neil Lareau[5], Craig Clements[5], and Marc L. Fischer[1]

[1]Lawrence Berkeley National Laboratory, Berkeley, California, USA, [2]California Institute of Technology, Pasadena, California, USA, [3]Department of Chemistry and Air Pollution Research Center, University of California, Riverside, California, USA, [4]ESRL, NOAA, Boulder, Colorado, USA, [5]Department of Meteorology and Climate Science, San Jose State University, San Jose, California, USA

Abstract Atmospheric inverse estimates of gas emissions depend on transport model predictions, hence driving a need to assess uncertainties in the transport model. In this study we assess the uncertainty in WRF-STILT (Weather Research and Forecasting and Stochastic Time-Inverted Lagrangian Transport) model predictions using a combination of meteorological and carbon monoxide (CO) measurements. WRF configurations were selected to minimize meteorological biases using meteorological measurements of winds and boundary layer depths from surface stations and radar wind profiler sites across California. We compare model predictions with CO measurements from four tower sites in California from June 2013 through May 2014 to assess the seasonal biases and random errors in predicted CO mixing ratios. In general, the seasonal mean biases in boundary layer wind speed ($< \sim 0.5$ m/s), direction ($< \sim 15°$), and boundary layer height ($< \sim 200$ m) were small. However, random errors were large (~1.5–3.0 m/s for wind speed, ~ 40–60° for wind direction, and ~ 300–500 m for boundary layer height). Regression analysis of predicted and measured CO yielded near-unity slopes (i.e., within 1.0 ± 0.20) for the majority of sites and seasons, though a subset of sites and seasons exhibit larger (~30%) uncertainty, particularly when weak winds combined with complex terrain in the South Central Valley of California. Looking across sites and seasons, these results suggest that WRF-STILT simulations are sufficient to estimate emissions of CO to up to 15% on annual time scales across California.

1. Introduction

A variety of methods have been applied to address the problem of accurately monitoring and mapping emissions of greenhouse gasses (GHGs). Bottom-up inventories have approached the problem by aggregating emission estimates from spatial and temporal distributions of known GHG sources to generate estimates of total emissions [e.g., *van Vuuren et al.*, 2009]. The bottom-up approach has a variety of strengths but can be limited by incomplete knowledge of processes that contribute to GHG emissions. On the other hand, atmospheric inverse modeling is a complementary technique that uses bottom-up emission estimates of GHGs in conjunction with meteorological Lagrangian transport models to estimate GHG emissions without requiring complete knowledge of individual sources.

In 2006 the state of California passed landmark legislation in Assembly Bill 32 (AB32), which committed the state to reducing GHG emissions to 1990 levels by 2020. This has led to focused efforts to measure, quantify, and mitigate emissions of a variety of key GHGs. One component of this effort was the establishment of multiple GHG monitoring towers throughout the state. In addition to efforts with regards to GHG mitigation, concerns about statewide air quality have contributed to monitoring, quantification, and mapping of sources of a variety of pollutants including carbon monoxide (CO). While CO has limited impact as a GHG, it does have specific advantages that make it an ideal candidate to assess the accuracy of atmospheric inverse modeling and develop estimates of potential biases and limitations in the modeling framework. The first advantage of CO is that there are few poorly constrained natural sources of emissions in California. Wildfire is a major natural source [*Pfister et al.*, 2005], and large wildfires are easily identifiable and of limited duration; these two properties are amenable to straightforward filtering of wildfire contributions to CO observations.

Additionally, total anthropogenic emissions of CO are tightly coupled to emissions from fuel combustion, which are inventoried for air quality control by the California Air Resources Board (CARB). These advantageous properties lend themselves to bottom-up estimates of CO emissions with low uncertainty relative to bottom-up estimates of emissions for other GHGs.

Lagrangian transport models track the transport of particles through the atmosphere. These models have been successfully used in a wide variety of applications. These applications include tracking intercontinental transport of pollution plumes [e.g., *Yienger et al*., 2000; *Fiore et al*., 2002], assessing atmospheric deposition of mercury [*Holloway et al*., 2012], and estimation of atmospheric dispersion of radioactivity [*Stohl et al*., 2012]. Within California, Lagrangian transport models have been used for tracking pollution and GHG emissions [e.g., *Brioude et al*., 2013; *Newman et al*., 2013]. Lagrangian transport models are invaluable for atmospheric inverse modeling of GHGs. In atmospheric inverse modeling they are used to estimate residence time footprints of air parcels reaching an observation point. Combined with estimates of surface emissions, these footprints provide estimations for the GHG concentration of the parcels. However, Lagrangian transport models are typically driven by simulated meteorology from regional atmospheric models or atmospheric fields from reanalysis products and are subject to uncertainties associated with deficiencies in these systems that must be quantified and accounted for in atmospheric inversions of GHGs [e.g., *Nehrkorn et al*., 2013; *Rogers et al*., 2013; *Angevine et al*., 2014].

In this study, boundary layer meteorological measurements, tower-based observations of CO mixing ratio, and bottom-up estimates of CO emissions across California were used to identify, quantify, and reduce biases and potential errors in the atmospheric inverse modeling framework. There were several specific objectives that this study addressed: (1) to determine Lagrangian transport model parameterizations by site and season in California that minimizes predicted CO error; (2) to separate and assess transport model random error and mean bias in boundary layer winds and boundary layer height; and (3) to accurately constrain estimates of uncertainty in GHG signal predictions due to a combination of factors including background mixing ratio estimation, random transport error, and observational uncertainty.

2. Methods

Following earlier emissions studies [*Zhao et al*., 2009; *Jeong et al*., 2012a, 2013], we used a Lagrangian transport model to estimate CO mixing ratios at five observation towers in California from June 2013 to May 2014. Two necessary components of this inversion technique were an atmospheric transport model and prior emission model, which were used to predict GHG mixing ratios at a target location at a specific time as

$$\mathbf{K} = \mathbf{FE} \tag{1}$$

where **K** is the vector of predicted mixing ratios, **F** is the linearized transport solution and each row of **F** is a footprint vector of length n (n is the number of grid points) describing the spatial distribution of residence times for particles tracked backward in time from the target location, and **E** is an a priori emission vector of length n providing the spatial distribution of estimated prior emissions using data from an emission model. Similar to *Jeong et al*. [2012a, 2013, 2016a, 2016b], this study used the coupled WRF-STILT (Weather Research and Forecasting and Stochastic Time-Inverted Lagrangian Transport) model to estimate atmospheric transport of GHGs [*Lin et al*., 2003; *Skamarock and Klemp*, 2008; *Nehrkorn et al*., 2010].

In this section we describe the observational data sets used in this study and provide details on the prior estimates of CO emissions in California, the atmospheric transport model, and the observations used to diagnose atmospheric transport errors. Additionally, the transport error analysis used to generate uncertainty statistics is described with a summary of how these statistics were propagated through the Lagrangian transport model. Finally, the methods used to compare measured and predicted CO signals are presented.

2.1. GHG and Meteorological Observation Sites

Greenhouse gas observations in California have been taken at 14 towers spread across California as shown in Figure 1 (see also supporting information Table 1). Among the 14 GHG sites, hourly mixing ratio measurements of CO were available during June 2013 to May 2014 at the CIT, SBC, and WGC sites, October 2013 to May 2014 for ARV, and daily measurements of CO were available June 2013 to May 2014 for STR (see Figure 1b). More detailed information of CO measurements can be found in Text S1 of the supporting information [*Andrews et al*., 2014]. Additionally, we used observations of boundary layer winds and heights from

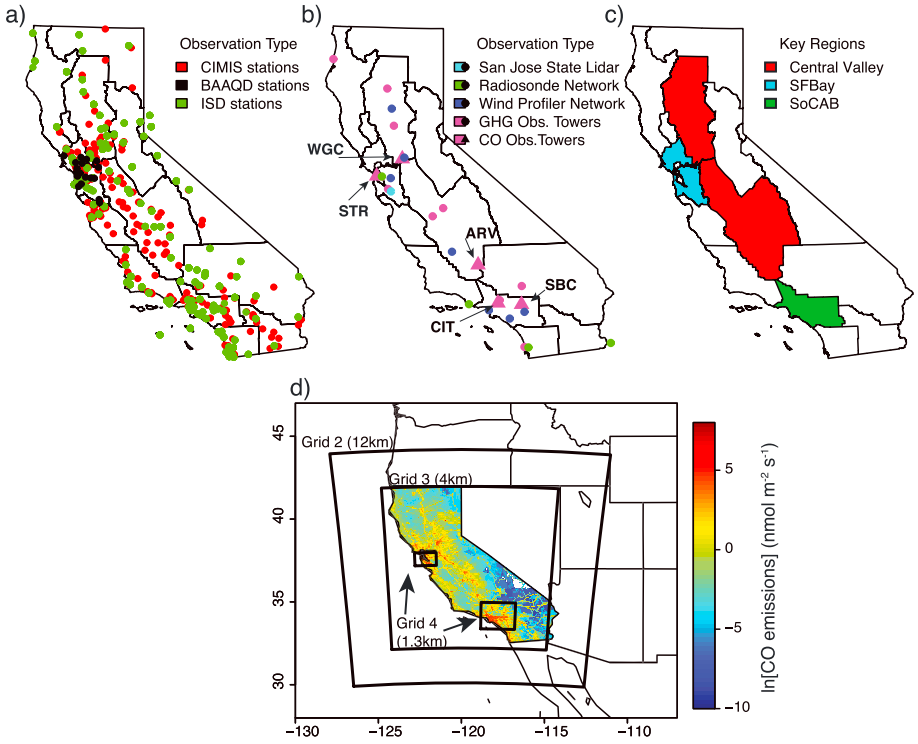

Figure 1. Locations of (a) meteorological stations; (b) tower observations (purple), radiosondes (green), and wind profilers (blue); and (c) key regions used in this study. Also shown are computational domains used in WRF-STILT and a priori (d) CO emissions from California. Note that an outer domain (36 km resolution) that included much of the western part of North America and the Northeast Pacific Ocean is not shown in this figure (grid extended from ~150°W–90°W to ~17°N–63°N).

four radiosonde locations, seven radar wind profiler stations, and a Doppler wind lidar to assess WRF meteorology and generate error statistics for propagation in STILT. Finally, wind measurements from surface stations across California and several GHG towers were used to test local WRF model meteorology and filter analysis. The locations of these measurements are shown in Figure 1a.

When available, wind profiler observations from July 2013 to May 2014 were used to assess errors in boundary layer height (z_i) and planetary boundary layer (PBL) winds (supporting information Table S1). These observations sampled winds throughout the boundary layer at an hourly time step. Estimates of daytime z_i were derived using a combination of automated algorithms and qualitative analysis from wind velocity observations of subhourly vertical velocity and returned signal strength [*Wyngaard and LeMone*, 1980; *Bianco and Wilczak*, 2002]. This was a manually intensive exercise that required subjective expert judgment [*Bianco et al.*, 2008]. As such, availability of wind profiler z_i data was limited relative to wind data, which were derived using an established automated algorithm. Additionally, the availability of wind profiler boundary layer

Table 1. Mean Seasonal CO Prediction Error (ppb) at ARV, CIT, SBC, and WGC Sites[a]

	ARV	CIT	SBC	WGC
JJA	NA	17.8 (9.7) [±6.6]	−22.0 (15.1) [±6.0]	20.8 (100.3) [±3.1]
SON	−22.6 (41.8) [±5.3]	22.4 (13.8) [±10.4]	−17.5 (12.0) [±11.2]	−12.5 (25.3) [±3.6]
DJF	−0.5 (1.1) [±5.0]	−9.7 (3.9) [±17.9]	−3.4 (2.8) [±8.9]	1.6 (2.7) [±5.6]
MAM	3.9 (17.8) [±2.0]	5.4 (3.8) [±7.9]	−16.4 (19.0) [±5.4]	8.7 (69.9) [±2.1]

[a]Values in parentheses (top row) represent the mean error as a percentage of the seasonal mean background-subtracted CO signal. The uncertainty of the mean error is also reported (ppb; in brackets; bottom row).

height observations was restricted to periods when a clear convective boundary layer exists. The fraction of days when wind profiler boundary layer heights were available varied by site and season. For example, at a wind profiler in Irvine, California, from September 2013 to April 2014, there was a minimum of a single missing day during August 2013 and a maximum of seven during January 2014. For days when the boundary layer height observations were available, there were between three to eight observations each day. Similar data frequency was available at other sites.

The wind profiler network in California has been in a continual state of flux, with profilers only operating for limited periods of time or experiencing extended downtime. In order to retain spatial coverage of profiler observations across California we accounted for wind profiler sites that were not operating from June 2013 to May 2014, or were missing months during that period, by filling in monthly data gaps with observations from prior years and prior corresponding WRF simulations from *Jeong et al*. [2012a, 2013]). We assumed that monthly random error statistics were approximately invariant across years but acknowledge that this was a potential source of error for assessing meteorological uncertainty in the GHG predictions. For locations with data during multiple years we found that the difference in error statistics including mean bias and random error magnitude across years to be significantly smaller than the difference of the same error statistics across sites. To focus on winds within or near the PBL, we limited our analysis of profiler winds to those observations that were below 2000 m above ground level. These winds directly influence the time atmospheric air parcels reside in the boundary layer over a given location, which is proportional to the influence surface emissions from a location have on estimated GHG signal. Finally, we removed outliers in PBL winds and z_i by removing those observations that were greater than three standard deviations from the mean daytime values.

At radiosonde sites within California, observations were generally taken one or two times a day. These observations extended throughout the troposphere, but as with wind profiler measurements above, we limited our analysis of wind speed and direction to the portion of observations below 2000 m. This is because the monthly-averaged diurnal convective boundary layer depths are lower than 2000 m in California [*Bianco et al*., 2011].

Observations of surface wind speed and direction were extracted from the Integrated Surface Database (ISD) [*Smith et al*., 2011], California Irrigation Management Information System (CIMIS; http://www.cimis.water.ca.gov/), and Bay Area Air Quality District databases (BAAQD; http://www.baaqmd.gov/). Although these measurements were limited to observing winds near the surface, they provided excellent temporal and spatial coverage across California, with hourly observations available at most sites (Figure 1a).

One element of uncertainty was determining the portion of model-observation error that should be assigned to observational inaccuracy. Based on error statistics from previous studies we assumed the magnitude of observational uncertainty for wind measurements to be $1\,m\,s^{-1}$ [*Strach et al*., 1987], the uncertainty of z_i measurements to be 200 m [*Dye et al*., 1995], and the uncertainty of CO measurements to be 3 ppb [*Andrews et al*., 2014].

2.2. Prior Emission Model Estimates and Background Atmospheric CO Mixing Ratios

To calculate model predictions of CO at the GHG observational sites, a priori estimates of surface emission rates of CO and atmospheric background mixing ratios of CO are required. We used hourly data provided by CARB and the European Commission Joint Research Centre and Netherlands Environmental Assessment Agency (Emission Database for Global Atmospheric Research (EDGAR)). The CARB data were provided hourly for 2012 with high spatial resolution (~4 km × 4 km), and coverage across California and its coastal waters, while EDGAR provided annual mean CO emission data for 2008 with global coverage at ~10 km resolution. We found the primary source of variation in CO emissions in California to be differences between weekdays and weekends. In order to retain these variations for the June 2013 to May 2014 period of interest we used the 2012 CARB data and a simple selection process. For each month from June 2013 to May 2014 we constructed a pseudo time series of California emissions by first determining the day of the week each day falls on. Then if the day was Monday-Friday we selected a random weekday from the corresponding month in 2012 to represent it. Likewise, if the day was Saturday or Sunday we selected a random weekend day from the corresponding month in 2012. Repeating this for all months from June 2013 to May 2014 ensured that all weekday emission estimates corresponded to weekday emissions in the original 2012 data and similar for weekends.

The annual mean map of prior emission estimates for California is shown in Figure 1d. Major metropolitan areas and highways are clearly evident. Outside of California emissions were set to values from the EDGAR data set which was interpolated to ~4 km resolution. This interpolation was done using a nearest-neighbor approach to retain sharp gradients that exist in the data set.

In a recent study focused on the Los Angeles Basin, *Brioude et al*. [2013] compared top-down inversion estimates of CO emissions from an intensive aircraft campaign with earlier bottom-up CARB estimates of surface emission rates for Southern California Air Basin (SoCAB). They found that CARB estimates were accurate to ~15% relative to their top-down estimates, with biases dependent on differences between weekday and weekend emissions. Based on this, we assumed that a priori emissions within each footprint calculation have a 15% uncertainty associated with them.

To estimate background atmospheric mixing ratio for CO, we followed the procedure of *Jeong et al*. [2012a]. A temporally and spatially varying mixing ratio "curtain" of CO was estimated at 130°W using vertical profiles of CO mixing ratios from aircraft along the West Coast U.S. and Hawaii, as well as marine boundary layer data from the Cooperative Air Sampling Network (http://www.esrl.noaa.gov/gmd/ccgg/flask.html). For background observations along the West Coast, particle trajectories were tracked backward in time to 130°W where they were binned by latitude (10° resolution) and altitude between 3000 and 7000 m (1000 m resolution). Below 1000 m, time- and space-varying values for the Pacific Ocean from the Cooperative Air Sampling Network were used. Bins between 1000 and 3000 m were calculated by vertical interpolation. Finally, the background CO signal was estimated for each time step by taking the mean background value across all particles. The background CO was added to the local signal calculated using footprints and prior emissions estimates to predict the total CO at each observational site for each hour from 1900–300 UTC.

2.3. The WRF-STILT Transport Model and Simulation Error Analysis

In the WRF-STILT model, Eulerian fields from WRF are used to drive Lagrangian particle transport in STILT. For this study we used a series of two-way nested grids in WRF to improve simulation accuracy in key regions of California (Figure 1d). The horizontal resolutions of these grids were 36, 12, 4, and 1.3 km, respectively. The outer 36 km grid simulated a region that included much of the western part of North America and the Northeast Pacific Ocean (not shown; grid extended from ~150°W–90°W and ~17°N–63°N). The 4 km grid encompassed the majority of California, and two 1.3 km grids were centered on regions surrounding the San Francisco Bay Area (SFBay) and the metropolitan area of Los Angeles in the Southern California Air Basin (SoCAB) to improve simulation accuracy in these key regions (Figures 1c and 1d).

Version 3.5.1 of the WRF model was used for all simulations. The boundary and initial conditions were extracted from the North American Regional Reanalysis data set (NARR) [*Mesinger et al*., 2006]. The primary WRF simulations were run from June 2013 to May 2014 using a series of 30 h runs that covered the entire simulation period. For each of the 30 h runs, data from the initial 6 h were considered to be spin-up and were discarded in the final analysis [*Pillai et al*., 2011; *Jeong et al*., 2013]. The remaining model output fields were saved hourly.

Due in part to extensive topography generating unique meteorological conditions in California, there have been multiple studies in recent years that have assessed the ability of a variety of WRF parameterizations to reproduce observed conditions in the region [*Bao et al*., 2008; *Michelson and Bao*, 2008; *Angevine et al*., 2012; *Jeong et al*., 2012b, 2013]. Here we largely follow the WRF setup described in *Jeong et al*. [2013]. We parameterized WRF to use 50 vertical levels to minimize errors in boundary layer meteorology over California's complex terrain. Radiative transfer used the Rapid Radiative Transfer model (RRTM) scheme for longwave radiation and the Goddard scheme for shortwave. The Purdue-Lin parameterization was used for microphysics and the Grell-Devenyi ensemble mass flux scheme for convection. Finally, as described in more detail below, we tested several combinations of parameterizations for the land surface and planetary boundary layer. For the land surface we used both the Community Land Surface Model (LSM) and Noah models and used the updated Mellor-Yamada-Nakanishi-Niino (MYNN2) and Yonsei University (YSU) parameterizations for boundary layer physics. The MYNN2 boundary layer parameterization was primarily used, but the YSU parameterization was selectively used to take advantage of an improved representation of topographic influences on boundary layer meteorology. The use of these WRF configurations was based on conclusions from *Jeong et al*. [2013]. In their study, they showed that transport modeling in California needs multiple

configurations, and footprints for towers in the Central Valley are distinct from those in Southern California. The purpose of the multiple configuration was to avoid mean biases in transport modeling. This is particularly prevalent in the Central Valley where lack of irrigation in certain land surface parameterizations likely leads to large overestimates of boundary layer height in the region.

To generate emission footprints, each day at each site 500 particles were initialized hourly from 1900 to 300 UTC (1100–1900 local standard time (LST)) in WRF-STILT. Following the work of *Jeong et al*. [2012a] the trajectories of these particles were calculated and tracked backward in time using WRF-STILT until either (a) the particles were transported outside the computational domain or (b) 7 days had passed. For each time step, when the trajectory of a given particle remained within the domain and within the atmospheric boundary layer, the time the particle spent over a given location and the local boundary layer height was calculated. By aggregating this information for all particles, footprints (**F** in equation (1)) were calculated [*Lin et al*., 2003]. During each hour that particles were released, the footprint strength at each grid point was combined with a priori estimates of surface CO emissions to predict local CO signals at observational locations due to emissions within the computational domain (equation (1)).

In addition to providing time for particles to move away from observational locations and sample remote regions, tracking the particles for up to 7 days provided sufficient time for a significant fraction of them to reach 130°W where they were assigned an observationally based CO mixing ratio representing the background CO for each particle using the Pacific CO curtain described in section 2.2.

Errors and limitations in the WRF representation of atmospheric processes can lead to transport errors in the STILT prediction of CO at observational sites. If these errors are not accounted for they can contribute to unrealistic posterior estimates of surface emissions from atmospheric inversions with unrepresentative uncertainty estimates. Two WRF variables that are of particular importance to our application of the coupled WRF-STILT included planetary boundary layer height (z_i) [*Gerbig et al*., 2008] and winds within and near the planetary boundary layer (PBL) [*Lin and Gerbig*, 2005]. The PBL winds influence both the path that STILT calculates Lagrangian particles take and the fraction of a particle's final CO concentration contributed by a given grid point. The fraction is directly proportional to the length of time a particle spends in the PBL over a given grid point (i.e., horizontal wind speed). Finally, z_i is inversely proportional to the atmospheric mixing ratio of boundary layer CO due to underlying surface emissions and, thus, the CO mixing ratio that Lagrangian particles in the PBL incorporate.

Due to the importance of z_i and PBL winds, errors in WRF's representation of these fields must be accounted for in the Lagrangian transport model. WRF-STILT has been constructed to account for transport errors in the *U* and *V* components of PBL winds by assuming that WRF errors (ε) have insignificant mean error and can be represented by a Gaussian centered at zero with standard deviation equal to the standard deviation of WRF error relative to observations ($\sigma_u(\varepsilon)$) [*Lin and Gerbig*, 2005]. Previously, errors were assimilated into the STILT framework by adding an error component (ε) that is representative of PBL wind uncertainty:

$$\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}' + \varepsilon \tag{2}$$

where **u** is the overall particle velocity vector, $\bar{\mathbf{u}}$ is the mean component, and $\mathbf{u}'$ is the turbulent component of particle velocity. ε was drawn from a Gaussian distribution with mean 0 and standard deviation equal to the observed standard deviation of boundary layer wind error ($\sigma_{u,v}$) and decorrelated exponentially as a particle moved in space or time based on horizontal (l_x), vertical (l_z), or temporal (l_t) length scales [*Lin and Gerbig*, 2005]. This formulation increased the spread of particles in STILT and allowed estimation of uncertainty from random WRF transport error, although systematic errors in wind speed and direction were assumed to be negligible. For regions where systematic errors exist this limits the accuracy of transport estimates and would be expected to lead to significant biases in CO signal estimates.

In this study, error statistics were generated from monthly time series. Application of equation (2) presented problems at urban sites that occasionally observed large CO concentrations due to stagnant conditions with very small wind velocities. Estimating transport error using equation (2) led to systematic biases during these periods, as slow particles with low wind speed had a high likelihood to increase in wind speed due to the addition of ε. This artificially moved particles outside of the local urban area and resulted in CO predictions systematically biased lower than control simulations during periods of low wind speeds and high CO predictions, and hence, unrealistic estimates of transport error that are assumed to be unbiased. To reduce this bias

we adopted the approach that *Gerbig et al*. [2008] introduced in STILT for accounting for z_i errors in WRF to use with boundary layer winds. Similar to *Lin and Gerbig* [2005], *Gerbig et al*. [2008] assumed a Gaussian error distribution for z_i and neglected the mean error component by assuming it to be small relative to random error, and the standard deviation of the random error Gaussian (σ_{zi}) was set equal to the standard deviation of WRF z_i errors relative to observations. However, *Gerbig et al*. [2008] scaled σ_{zi} by the mean observed z_i and then introduced as a multiplicative factor directly influencing particle mixing ratio, instead of an additive term as in equation (2). We adapted this for estimating transport error by boundary layer winds by scaling $\sigma_{u,v}$ by the monthly mean wind speed for each site and applying this value as a multiplicative factor on particle velocity in both the *U* and *V* directions. This procedure removed the artificial CO bias at low wind speeds. Finally, following the approach of *Lin and Gerbig* [2005] we assessed the fraction of uncertainty in CO prediction that can be attributed to simulated random meteorological errors by calculating the mean difference in variance of the 500 particle CO mixing ratios between a control run with neither boundary layer nor wind error applied and separate simulations with uncertainty in boundary layer height and wind velocity included.

2.4. Regression Methodology and Data Filtering for CO Analysis

WRF-STILT is a state-of-the-art atmospheric transport model that has been successfully used to estimate surface emissions in a variety of studies. Additionally, the CARB estimates of surface emissions of CO potentially represent the most accurate and detailed spatially explicit bottom-up prior estimate of surface emissions available for any gas [*Brioude et al*., 2013]. However, limitations remain that required periods to be filtered from the analysis. In this study, the key limitations included meteorological conditions that violated our assumption that particle trajectories originate from the Pacific Ocean and sample background mixing ratios of CO at 130 W, the representation of wildfires in a priori surface emissions, and limitations due to periods of WRF meteorology and transport deviating significantly from observed conditions.

To reduce the impact of particle trajectories not reaching 130°W, we excluded CO predictions during periods when low fractions of the particles' trajectories reached 130°W. During October 2013 to May 2014 we removed results from hours when this fraction was less than 0.8. For June 2013 to September 2013 we relaxed this slightly to 0.7 in order to retain significantly more CO prediction periods due to seasonal shifts in meteorology reducing the fraction of particles reaching the 130°W CO mixing ratio curtain.

Wildfires represent very distinct periods of emissions from relatively small areas. Moreover, the location and timing of wildfires changes each year. Since the CARB CO prior emission inventory used in this study was developed for 2012, emissions from wildfires would not be valid for the 2013–2014 period of this study and were not used. To mitigate the influence of wildfires on our comparison of CO predictions and observations, we excluded periods when observations appeared to have been influenced by wildfires. The location and duration of major wildfires were identified using the CALFIRE database for 2013 and 2014 (http://www.fire.ca.gov/general/firemaps.php). Hourly footprints for each site were inspected to determine the likelihood that an identified fire could be influencing CO observations. Periods when it was likely that wildfires were influencing observations were excluded from the analysis. However, it should be noted that this analysis was limited to relatively large fires included in the CALFIRE database, and the potential impact of small-localized burns may not have been accounted for.

Regional WRF meteorological errors were identified using observations of wind speed and direction from ISD, CIMIS, and BAAQD surface stations near GHG observation towers (within 50 km). Periods when the wind speed error normalized by observed wind speed exceeded 0.5 or the mean difference in wind direction between observed and simulated surface winds exceeded 90° were filtered from the analysis. Additionally, at the ARV GHG measurement site we identified model deficiencies in transport for southeasterly winds (70–190° east of north) from November 2013 to January 2014, which led to periods of large underestimates in predicted CO that persisted across all tested boundary layer—land surface model combinations. These periods were filtered from the CO analysis. We hypothesize that this deficiency may be due to excessive boundary layer venting of air parcels during these periods due to difficulties in the representation of complex local topography.

To account for uncertainty in both predicted and observed CO concentration values, we used Type II regression (geometric linear regression) for all regression estimates in this study. This regression is valid given that correlation between predictor and predictand is significant. In this study, correlation coefficients

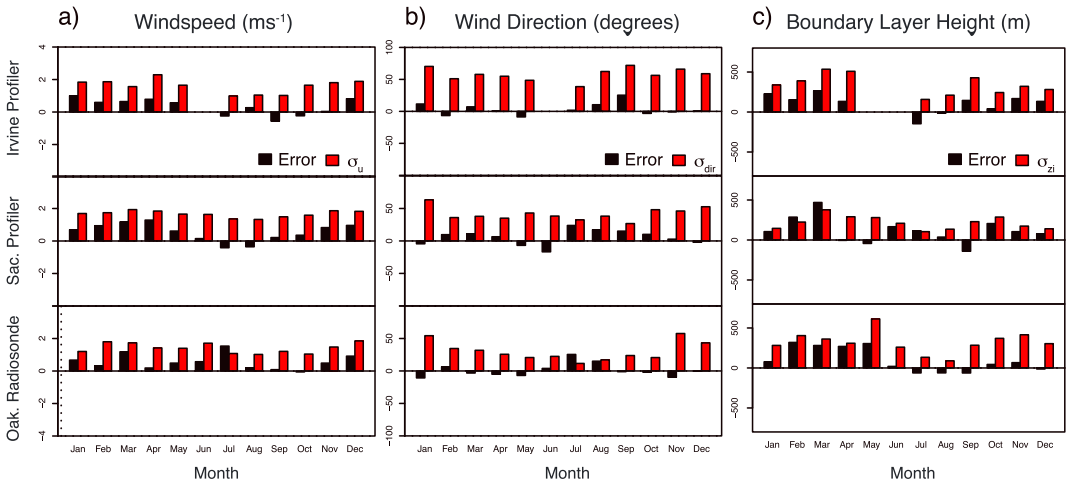

Figure 2. Monthly mean (black bars) and random (red bars) error for (a) wind speed (m s^{-1}), (b) wind direction (degrees), and afternoon (1200–1700 LST) (c) boundary layer height (m) for Irvine, Sacramento, and Oakland stations. Wind speed and wind direction errors are calculated from 0 to 2000 m.

for all regressions shown were found to be statistically significant ($p <= .01$). In this regression the sum of the squared perpendicular deviations is minimized to estimate the regression slope and associated slope uncertainty.

3. Results

3.1. Evaluation of Meteorology Errors

There are several regions of California that are of particular importance in terms of magnitude of GHG emissions. These include California's largest urban areas of the San Francisco Bay Area (SFBay) and the Southern California air basin (SoCAB; which encompasses Los Angeles). Also, extensive agriculture in the Central Valley (CV) has been shown to generate significant emissions of GHGs associated with livestock and fertilizer (e.g., CH_4, NO_2) [*Jeong et al.*, 2012b, 2013]. Therefore, it is of particular importance to model these regions accurately and account for errors in boundary layer meteorology. The first column of Figure 2a shows 0–2000 m daytime mean monthly values of $\sigma_{u,}$ and the monthly mean errors of WRF wind speed relative to observations using data from the Sacramento profiler in the CV, the Irvine profiler in SoCAB, and the Oakland radiosonde station in the SFBay (see Figure 1 for the wind profiler sites associated with individual GHG sites). We note that some sites (e.g., ARV) may not be well represented by the closest available wind profiler, potentially resulting in undiagnosed transport errors. For this reason, we compare measured and predicted CO signals as a further diagnostic to help quantify errors and identify potential transport errors due to mean biases in meteorology. These stations were selected based on their location and data availability from June 2013 to May 2014. Supporting information Table S1 presents the seasonal mean and random error (represented by the standard deviation of model-observation error) of wind speed, wind direction, and boundary layer height representative of each GHG site. The error statistics in supporting information Table S1 were calculated using data from WRF-STILT meteorology and observations from profiler or radiosonde sites with available data nearest to the associated GHG site. There were several inferences that could be made from Figure 2. First, the magnitude of the random error (σ_u) was considerably larger than the magnitude of the mean error. The only exception to this was July in SFBay, which was likely due to limited numbers of daytime radiosondes launched (~1 daytime launch each day) during relative to the continuous profiler measurements, and WRF not properly capturing a synoptic event early in the month. Additionally, the random error, σ_u, tended to be at a yearly minimum during the summer. This was due to relatively strong and stable zonal winds during these months. Finally, inspection of histograms confirmed that the wind errors were approximately Gaussian about the mean.

One limitation of using wind speed as an assessment of meteorology is that its scalar nature does not account for errors in wind direction. Errors in WRF wind direction can lead to erroneous STILT particle trajectories and

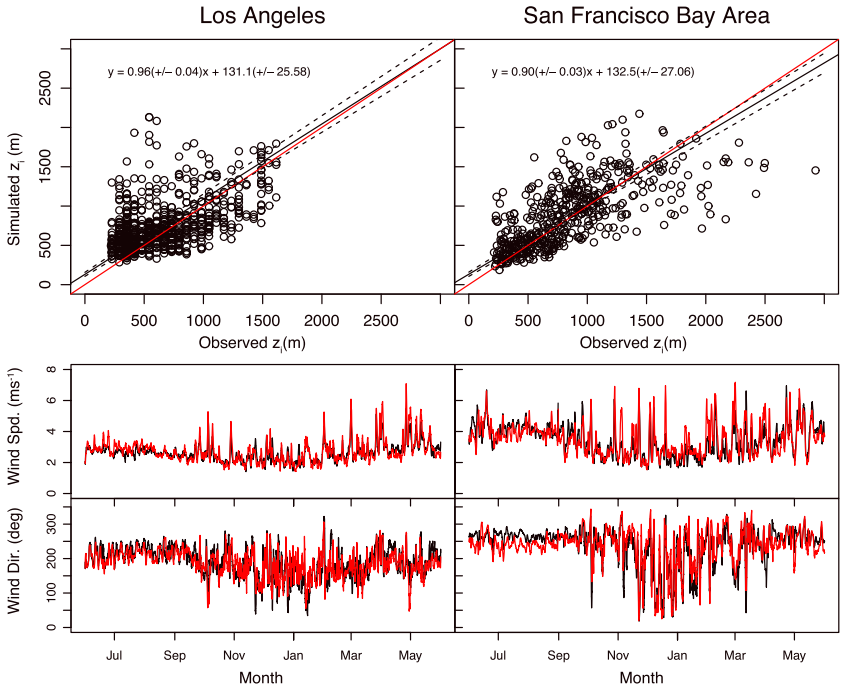

Figure 3. Simulated and observed z_i (1200–1700 LST) scatterplot for (top left) Los Angeles and (top right) San Francisco Bay area using the Irvine profiler and San Jose State University (SJSU) lidar, respectively. z_is from Irvine were available intermittently from June 2013 to April 2014. SJSU z_i data were available from December 2013 to May 2014. The red line indicates the 1:1 line. June 2013 to May 2014 time series of simulated (red line) and observed (black line) surface wind speed and direction at surface stations within 50 km of (bottom left) CIT and (bottom right) STR.

hence unrealistic emission footprints that may not properly sample prior emission estimates in the presence of sharp gradients in GHG emissions (such as near edges of urban areas or along major highways). The second column of Figure 2b shows the mean directional error (in degrees) and the standard deviation of wind direction error (σ_{dir}) for the three stations described above. Unlike the CV and SFBay area sites, σ_{dir} at the SoCAB site did not exhibit clear seasonality. Instead, it had consistent values of 60–80°, whereas σ_{dir} of the CV and SFBay sites ranged from ~50 to 60° during the winter to ~10–30° during the summer. At all sites the mean directional error was small relative to the random error.

For z_i the mean error was comparable in magnitude to the random error (Figure 2, third column), and in some cases such as February and March 2014 for the Sacramento profiler, the mean error exceeded σ_{zi}. Also, similar to the model-data errors in wind speed discussed above, errors in z_i did have some seasonality associated with them. February–May commonly exhibited the greatest mean and random errors across observational sites, while during July–August mean and random errors in z_i were at a minimum. However, due to the relatively low number of boundary layer height observations available, the significance of this difference was difficult to assess. Seasonally similar results were obtained at other sites as well (supporting information Table S1). For our study, this brings into question the validity of the assumption in *Gerbig et al*. [2008] that the mean error can be neglected and is a potential source of error in our transport model. As such, errors that arise from mean transport model bias are reflected in CO signal calculations and contribute to our estimates of overall transport error through their influence on regression slopes of observed and modeled CO concentration and differences in mean concentrations.

CO emissions are dominated by the major urban centers in SoCAB and SFBay, and agriculturally dominated portions of the state such as the Central Valley emit relatively small amounts of CO. Therefore, ensuring that the WRF-STILT model is accurately simulating meteorological conditions in these regions is vital. Figure 3 shows June 2013 to May 2014 time series of surface wind speed and direction for SoCAB and SFBay, and scatterplots of observed versus simulated boundary layer heights using WRF-STILT parameterized with MYNN2-Noah. For both SoCAB and SFBay, regression analysis indicates slopes of simulated versus

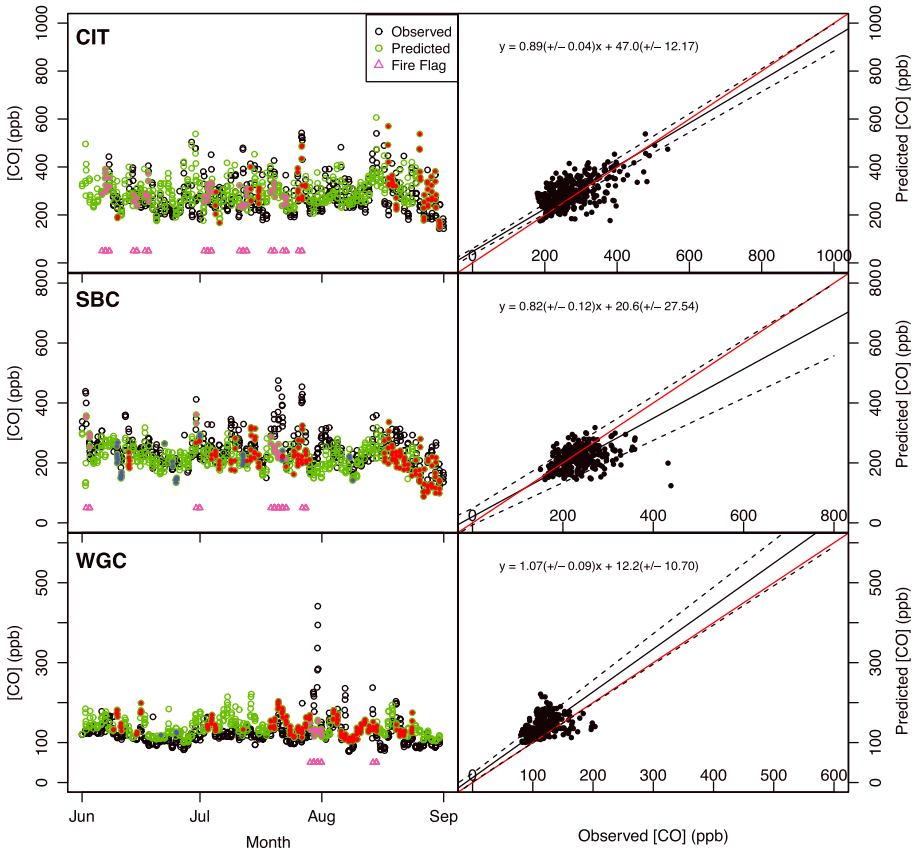

Figure 4. June 2013 to August 2013 time series and regression for (top row) CIT, (middle row) SBC, and (bottom row) WGC. Green circles represent predicted CO mixing ratios, and black circles represent observations. Filled circles indicate that the data were excluded from the regression analysis due to insufficient particle trajectories reaching 130°W (red fill), poor WRF representation of meteorology (blue fill), or wildfire (magenta fill and triangles).

observed z_i are within 1 of unity (0.94 ± 0.04 for SoCAB; 0.90 ± 0.03 for SFBay) with relatively small positive intercepts ~130 m. Additionally, inspection of simulated wind speed and direction relative to surface station observations revealed that WRF-STILT was capturing both seasonal changes and synoptic shifts in surface winds. Combined, these results give confidence that WRF-STILT was sufficiently simulating meteorological transport of emissions in key regions.

3.2. Seasonal WRF-STILT CO Predictions and Optimization of WRF Parameterization

At ARV, CIT, SBC, and WGC we assessed the accuracy by which a variety of parameterizations for WRF-STILT reproduced the observed hourly CO mixing ratios, surface winds, and boundary layer heights (where available) for June 2013 to May 2014. We found that no single model configuration was sufficient to optimally capture the CO dynamics across California due to extensive irrigation in the Central Valley and complex topography. Instead, we optimized WRF-STILT parameterization on a site-by-site basis. This is described in more detail in the supporting information (Text S2) [*Kueppers et al.*, 2007; *Sacks et al.*, 2009; *Sorooshian et al.*, 2011; *Harding and Snyder*, 2012; *Bagley and Miller*, 2015]. The specific WRF-STILT land surface and boundary layer parameterizations used at each CO site in this study are shown in supporting information Table S2.

Using the model configurations in supporting information Table S2, we seasonally analyzed the predicted CO signal at ARV, CIT, SBC, and WGC for summer 2013 to spring 2014. We used a seasonal time scale to minimize the impacts of potential modeling biases of individual synoptic meteorological events that can negatively influence CO predictions for several days and dominate error statistics over shorter time periods. This was particularly relevant for CO as its sources were primarily constrained to small regions of large anthropogenic influence such as cities and roads. Therefore, small errors in footprint estimation due to short-term

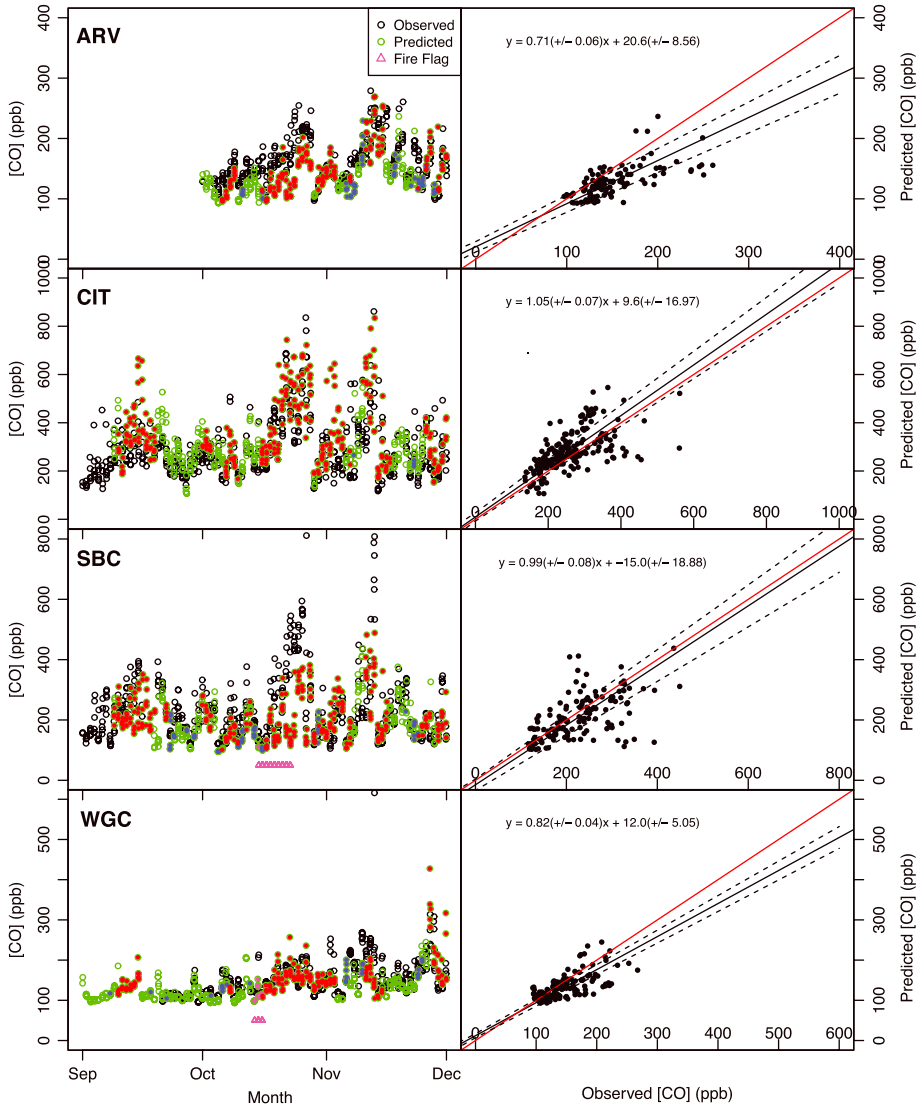

Figure 5. Same as Figure 4 but for September 2013 to November 2014. Additionally, (top row) ARV is included.

meteorological errors can lead to errant estimation of emission sources contributing to observed CO at a tower. The mean error of WRF-STILT predicted CO normalized by mean observed CO is summarized for all sites and seasons in Table 1.

At CIT, SBC, and WGC during June–August of 2013 (ARV CO observations were unavailable during this period) the predicted and observed CO signals (Figure 4) were small relative to fall and winter (Figures 5 and 6). In June and July there were multiple periods of potential wildfire influence at the southernmost sites (SBC and CIT), which were excluded from the analysis (magenta triangles in Figures 4–7). However, with the exception of July and late August for SBC, only small fractions of these months failed to satisfy our criteria for parcels sampling background conditions at 130°W (blue-filled circles). In the periods when predictions of CO were retained (open green circles in Figures 4–7), the magnitudes of mean error of predicted CO at CIT, SBC, and WGC were 10–22 ppb (Table 1). However, the sign of the errors in CO were not consistent across sites, with simulations overpredicting CO at CIT and WGC and underpredicting CO at SBC. As shown below, the underprediction of CO at SBC was consistent across all seasons but the range of both measured and predicted signals were smaller during summer. At CIT and SBC, removing background CO and normalizing the prediction error of CO by the mean observed CO led to errors in predicted CO between 9 and 15% of the background-subtracted observed CO during summer of 2013 m (Table 1). However at WGC, the

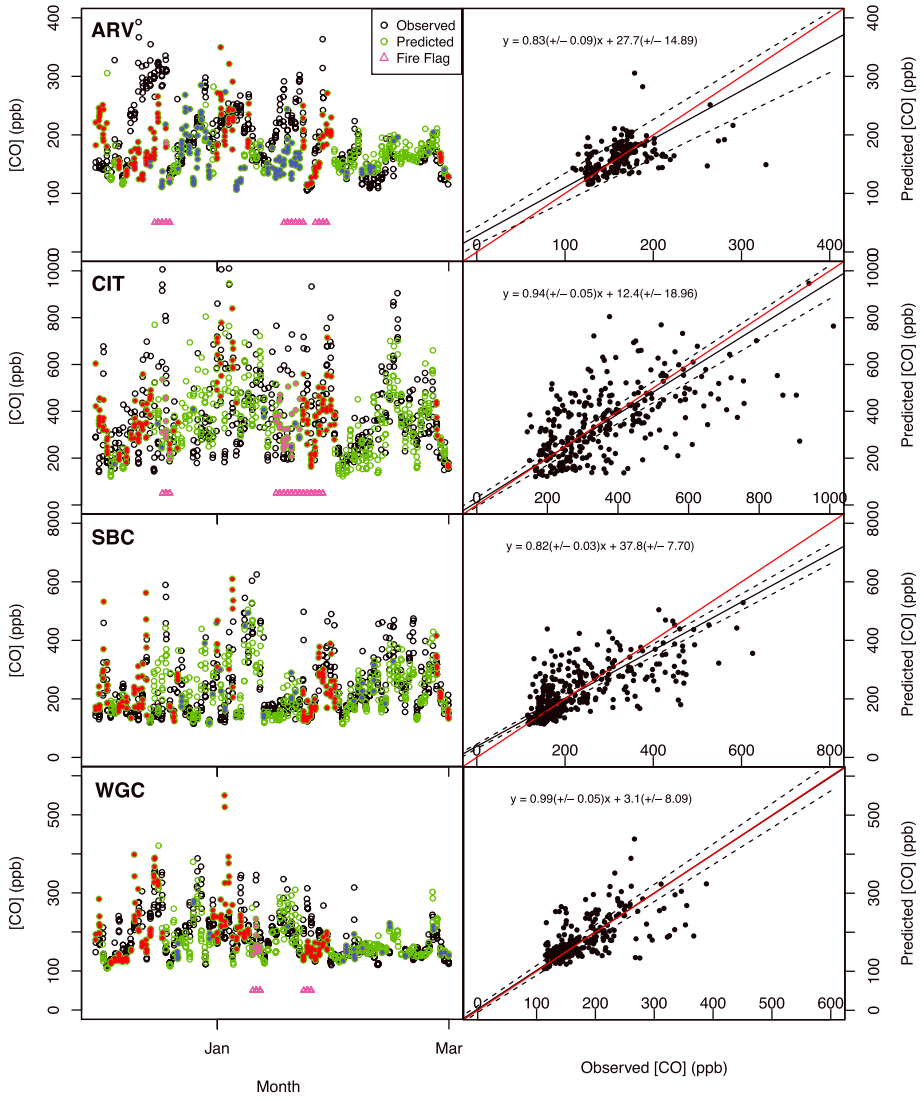

Figure 6. Same as Figure 5 but for December 2013 to February 2014. Note that blue-filled circles in time series also include periods of southeasterly winds at ARV.

fractional prediction error was much larger. Strong zonal winds and tall boundary layer heights contributed to CO signals that were typically small (~20 ppb) compared to other sites. Predicted CO was typically larger, leading to a mean overprediction in CO of 21 ppb. Combined with the consistent underprediction of CO at SBC, this may suggest that consistent undiagnosed mean transport model bias is influencing CO prediction for these sites and seasons. Nevertheless, regression analyses showed that overall the predicted and observed CO were generally consistent with observations across all sites, with slopes close to 1 (within 0.2) with values of 0.82 ± 0.12 at SBC, 0.89 ± 0.04 at CIT, and 1.07 ± 0.09 at WGC. At WGC and CIT, regression slopes near unity, small mean bias, and small slope uncertainty indicated that WRF-STILT accurately captured signal of CO for these sites during summer of 2013. At SBC, the combination of small mean error in CO, relatively large uncertainty of the regression slope, suggests that while WRF-STILT reproduced the mean seasonal CO, it did not capture the small summertime variations in CO.

Figure 5 compares the predicted and observed CO signals at ARV, CIT, SBC, and WGC for September-October-November (SON) of 2013. During the fall, a combination of changes in wind direction, intermittently low wind speeds, and small daytime boundary layer depths produced mean CO signals significantly larger than those found during the summer of 2013. However, the periods of highest CO also tended to be during periods

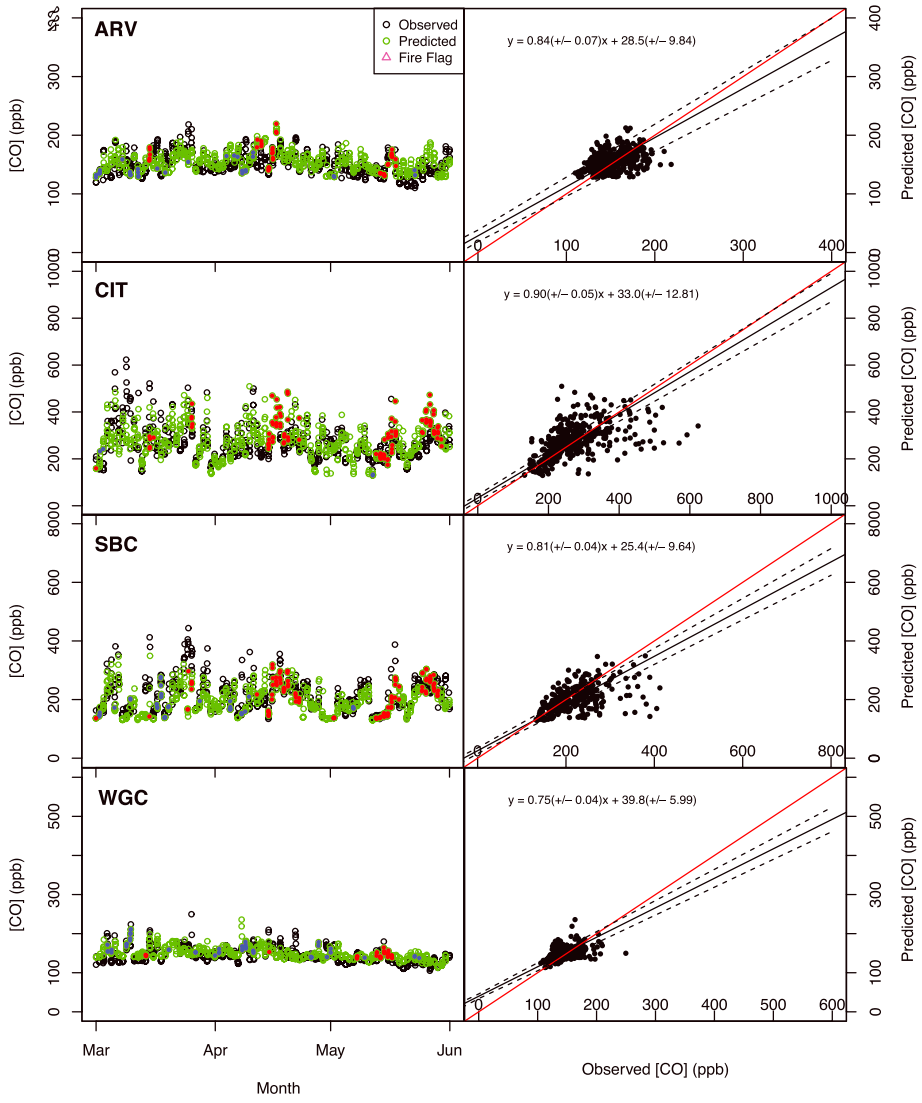

Figure 7. Same as Figure 5 but for March 2014 to May 2014.

when the background sampling criteria were not met, typically due to either very stagnant air masses over California or meteorological conditions transporting continental instead of Pacific air to the GHG towers. This was particularly prevalent at CIT and SBC. Wildfires were not factors during the fall at any sites, with the exception of short periods at SBC and WGC during October. Analysis of regression between observed and predicted CO revealed strong relationships during SON of 2013. At CIT and SBC the regression slopes were close to one with small uncertainty (1.05 ± 0.07 for CIT; 0.99 ± 0.08 for SBC). However, at WGC and ARV the slopes and mean prediction errors (Table 1) indicated WRF-STILT underpredicted CO mixing ratios at these sites. This was particularly noticeable at ARV where we speculate that intermittent flows and stagnant conditions in the complex topography at the southern end of the San Joaquin valley may have contributed to a mean prediction error of CO, which was 23 ppb lower than observed and the slope of the regression line was 0.71 (±0.06).

Similar to fall, the winter months of December 2013 to February 2014 (DJF) had large predicted and observed CO relative to spring and summer (Figure 6). Here, however, mean prediction errors were small at all sites, ranging from −9.7 at CIT to 1.6 ppb at WGC. When normalized by the mean background-subtracted observed CO, these represented mean errors of only 1–4%. This suggests mean transport errors were likely small during this season. However, we note a relatively large slope uncertainty (0.81 ± 0.09) and intercept uncertainty

(27.7 ppb ± 14.89 ppb) at ARV. As discussed in section 2, we excluded periods of southeasterly winds from the Mojave desert at ARV between November 2013 and January 2014 that coincided with significant underprediction of CO for all model configurations. This exclusion occurred during a significant fraction of the winter season (blue-filled circles in Figure 6). The source of this underprediction remains undiagnosed as errors in regional surface wind speed and direction were not obviously biased during these periods. However, these periods did typically correspond to periods of slow moving stagnant air in the region. This requires further analysis, and it is possible that this issue may be influencing periods that were not identified by our crude exclusion of southeasterly winds from November to January at ARV. SBC exhibited a similarly low regression slope (0.82 ± 0.03), but this result did not include the extensive exclusion of data as at ARV. At CIT, SBC, and WGC this was not an issue as evidenced by inspection of the time series, small mean error in CO predictions, and regression slopes close to one with small uncertainties (~ ± 0.05).

During spring (March–May) tall boundary layer heights and strong zonal winds contributed to small CO signals at all sites, (Figure 7). Additionally, shifts in prevailing winds and the absence of wildfires contributed to only a small fraction of the CO predictions requiring exclusion. With the exception of SBC, seasonal mean CO was overpredicted by 4–9 ppb across all sites. At SBC, seasonal mean CO was again underpredicted (−16.4 ppb). Also, at all sites the regression slopes were less than 1 and ranged from 0.81 at SBC to 0.94 at WGC. However, as discussed in the following section, small CO signals at ARV and WGC during these months may have led to small errors in background estimates of CO exerting relatively large influences on regression results.

Although problematic periods remain, the combination of seasonal biases in predicted CO typically smaller than 15 ppb and slopes of regression lines within 0.20 of unity for most sites and seasons (in most cases consistent within the uncertainty) lends confidence to WRF-STILT's ability to accurately predict CO and hence other GHGs at our observational towers. Interpretation of these results requires mention of some sources of potential uncertainty that may influence predictions of CO including random error in the transport model, uncertainty in the background CO concentration curtain, and limitations in the bottom-up CO emission model.

3.3. Background and Random Meteorological Error Analysis

We estimated the potential impact of uncertainty in background estimates of CO using observations and WRF-STILT CO predictions at the STR tower site (Figure 1b). STR is a tall (sample stream intake at 232 m) coastal site in San Francisco, and its observational footprint commonly samples relatively clean oceanic air, with minimal local influence on observed CO. As a result, STR was used to estimate potential uncertainty in background CO. We restricted the observational periods used for this analysis to those times when two requirements were met: (1) The surface wind measured at several coastal and offshore buoy stations near STR had to be consistently westerly for 8 h preceding an observation, and (2) the difference between the total simulated CO and background CO at STR was required be less than 5 ppb (indicating minimal continental influence). Unlike ARV, CIT, SBC, and WGC where hourly CO data exist, CO observations were only taken once a day at STR for most days from June 2013 to May 2014 (October 2013 excluded). After filtering the data using the above requirements, 35% of the available observations were retained (74 individual days), with the majority of the days retained being during nonwinter months due to seasonal shifts in prevailing wind direction.

Comparing the filtered STR CO observational data with corresponding WRF-STILT simulated CO revealed a mean error in simulated background CO of 4.5 ppb ± 1.4 ppb, with a root-mean-square error (RMSE) of 13.1 ppb. At sites during seasons with small observed mixing ratios of CO, such as WGC during June-July-August (JJA), and both ARV and WGC during March-April-May (MAM) the estimated RMSE of background CO was comparable to the total RMSE of simulated CO (31.8 ppb at WGC for JJA; 19.4 ppb at ARV for MAM; 20.2 ppb at WGC for MAM). Although the mean error in background CO was less than 5 ppb, the relatively large RMSE indicated that a large fraction of the random CO error at sites during seasons with small observed CO may be attributed to uncertainty in background CO mixing ratios.

Next, we examined the random component of uncertainty in predicted CO. Table 2 presents the RMSE of CO as a percentage of the mean CO signal associated with propagating random wind and boundary layer height errors using the error statistics given in supporting information Table S1. We found that the random errors associated with uncertainty in wind and boundary layer height ranged from 0 to 70 ppb depending on the

Table 2. CO Mixing Ratio Uncertainty Due To Random Error of Boundary Layer Height and Boundary Layer Winds Added in Quadrature (in ppb), and Uncertainty as a Percentage of Seasonal RMSE of CO Prediction (in Parentheses)[a]

	ARV	CIT	SBC	WGC
JJA	NA	16.1 (28)	9.6 (18)	0.3 (1)
		[0.37]	[0.32]	[0.43]
SON	3.2 (9)	28.1 (39)	8.4 (12)	1.7 (6)
	[0.39]	[0.43]	[0.23]	[0.51]
DJF	25.3 (77)	69.4 (50)	35.6 (47)	12.6 (29)
	[0.87]	[0.51]	[0.50]	[0.43]
MAM	2.9 (15)	33.6 (47)	10.9 (22)	1.9 (9)
	[0.71]	[0.51]	[0.35]	[0.67]

[a]Also shown is the CO uncertainty due to the combined uncertainty in boundary layer height, wind speed, observational uncertainty (assumed 3 ppb), and background uncertainty (~13.1 ppb) as a fraction of the seasonal RMSE for predicted CO (in brackets; bottom row).

site and season. These errors corresponded to 0–77% of the observed RMSE with mean biases removed (Table 2). The largest errors associated with meteorological uncertainties were during the winter months. Additionally, the largest fraction of error associated with random meteorological errors was almost always at CIT. (30–50% of RMSE). Since the influence of imposed random z_i and wind errors reduces with distance from the tower sites, footprints that are influenced by remote regions will be less influenced by local uncertainty in meteorology and associated strongly heterogeneous urban emissions. It is also interesting to note the large influence of meteorological errors on ARV during DJF (77% of total RMSE). This strong sensitivity during DJF may provide some insight into the difficulties of simulating CO at this site during winter months. Finally, the influence of random meteorological uncertainty at WGC is minimal during summer and fall. This was likely due to strong zonal winds during these seasons causing the SFBay region to be the primary influence on CO at WGC. Since the SFBay region is sufficiently distant from WGC and the CO signal has small variability, the application of local random errors on the effective footprint of WGC is small.

When we combined the random meteorological uncertainty with uncertainty associated with estimated background CO (RMSE = 13.1 ppb) and estimated observational uncertainty (assumed to be 3 ppb), these uncertainties accounted for between 39 and 87% of the total RMSE at ARV and WGC but only 23–51% of the total RMSE at CIT and SBC (Table 2). The remaining uncertainty was likely due to a combination of undiagnosed systematic deficiencies in WRF-STILT transport that led to mean biases in CO prediction, limitations in our assumption that a priori CO emissions have minimal error (including short-term variability) within the tower footprints, and errors associated with unresolved subgrid scale processes (aggregation error). There is also a possibility that the error statistics derived for each GHG site were not sufficiently representative of meteorological uncertainty. This could be due to error statistics for each GHG site being calculated from atmospheric profiler or radiosonde measurements that may be a significant distance from the GHG site (Figure 1b) and may be derived from prior years for cases where observations were not available for 2013–2014.

4. Discussion

We have assessed the magnitude of meteorological errors in WRF by region and season using the most comprehensive set of surface and boundary layer observations available, estimated CO bias and random error at four sampling sites in California, and constrained the influence of errors in background signal and transport by propagating observed random errors in WRF-STILT. This was motivated by the assumption that error in a priori estimates of CO emissions was small relative to other GHGs, such as CO_2, CH_4, and N_2O, and could hence be used to constrain the influence of background and transport model error on predicted emissions that encompass key regions of California. In this section we discuss the implications of this work for atmospheric inverse modeling of other GHGs in California and the limitations of our results.

First, our analysis of boundary layer meteorology for GHG sites across California showed that the mean errors associated with wind speed, wind direction, and boundary layer height were small (~0.5 m/s, < ~15°, and < ~200 m, respectively). Additionally, the small mean biases in predicted CO concentration suggest

that mean biases in transport meteorology are small or that mean transport biases are offset by errors in prior emissions or background concentration estimates. There was also some seasonality and spatial variability in the results that should be taken into account in studies of other GHGs. In particular, random errors in boundary layer meteorology tended to be minimized during summer months and largest during the winter when intermittent synoptic events were dominant.

We found that the optimal parameterizations for key WRF-STILT boundary layer and land surface representations changed by season and site across California. In particular, changes in these parameterizations were necessary to capture influences of complex topography at southern California sites and address the impact of irrigation on boundary layer heights in irrigated agricultural regions like the Central Valley. This suggests that WRF-STILT parameterizations that have been optimized for one location or season should not be assumed to be directly transferrable to other sites in very different locations and still expect best results. It also suggests that incorporating irrigation into all WRF-STILT configurations should be a priority for studies of regions where irrigation is prevalent.

Filtering transport model predictions based on surface meteorology of wind speed and direction, in addition to standard fire and background trajectory filtering, reduced transport model error and improved prediction of CO concentrations at the GHG towers. This is not currently a standard practice for transport model application. However, it is unreasonable to expect accurately modeled footprints for periods with significant errors in modeled meteorology, which can contribute to undiagnosed errors in GHG inversions. This filtering methodology represents a straightforward way to reduce the major inconsistencies in local meteorology and improve transport model predictions.

Using the WRF-STILT parameterizations that best represented local meteorology, we found the mean bias of predicted CO relative to observations to be within 15 ppb for most sites and seasons across California and observed CO-simulated CO regression slopes to be near unity (1.0 ± 0.20). This indicated that WRF-STILT was successful in both generating predicted CO mixing ratios with small biases and capturing the dynamics necessary to simulate CO observations that can change on diurnal, synoptic, and seasonal time frames. This gives confidence in the capability for WRF-STILT to simulate the dynamics of other GHGs for most seasons and sites across California with minimal mean biases. This suggests that either mean biases in transport meteorology are small or are offset by deficiencies in prior emissions or background concentration estimates. However, there are periods when important uncertainties remain and need to be treated with care. In particular, the winter months seem to present problems at several sites (i.e. ARV and SBC) that require additional analysis to determine if deficiencies at these sites were due to limitations in a priori emission estimates or if improved model physics are needed to accurately capture the observed GHG dynamics in the absence of extensive data filtering. Additionally, the development of more sites with sufficiently precise CO measurements would be useful for constraining and evaluating errors in transport and a priori emission estimates.

We have constrained the potential influence of random errors in background CO, observational error, and uncertainty in meteorology at observational sites. Overall, the observational error was found to be minimal relative to the uncertainty of background CO and meteorology, and combined these sources of uncertainty were between 35 and 85% of the RMSE. Background uncertainty was dominant at sites with small seasonal CO signals such as WGC during spring and summer, and ARV during spring and fall. Meteorological uncertainties were dominant at CIT, likely due to its urban location, and were largest during the winter at all sites. ARV was shown to be particularly sensitive to meteorological uncertainty during the winter, which may provides some guidance for improving the season's CO prediction.

Finally, we provide an approximate estimate of the likely mean bias in inverse model estimate of CO emissions across California due to systematic terrors in predicted transport variables. Here we observe that atmospheric inversions typically multiplicatively scale prior emissions for multiple areas (air basins or pixels) so as minimize the weighted difference between predicted signals and background-subtracted observations [e.g., *Jeong et al.*, 2013]. To first order, this amounts to optimizing the slope of predicted versus measured mixing ratios, and hence, the multiplicative error of an atmospheric inversion will be roughly equivalent to the deviation from unity of the regression slope between observed and predicted GHGs. For example, at CIT the deviation of regression slopes from unity range from 0.05 to 0.11, indicating potential uncertainty in the atmospheric inversion of GHGs at this location of ~10%. We note that while we do not have an independent estimate of systematic error in regional CO emissions, we might expect that the mean bias in

Table 3. Combined SoCAB and Central California Mean Regression Slope Bias and Associated Uncertainty[a]

	Mean Regression Slope Bias and Slope Uncertainty		Seasonal Slope and Uncertainty for Individual Sites			
	SoCAB	Central CA	ARV	CIT	SBC	WGC
JJA	-0.12 ± 0.04	0.07 ± 0.09	NA	0.89 ± 0.04	0.82 ± 0.12	1.07 ± 0.09
SON	0.02 ± 0.05	-0.21 ± 0.03	0.71 ± 0.06	1.05 ± 0.07	0.99 ± 0.08	0.99 ± 0.04
DJF	-0.15 ± 0.03	-0.05 ± 0.04	0.83 ± 0.09	0.94 ± 0.05	0.82 ± 0.03	0.99 ± 0.05
MAM	-0.15 ± 0.03	-0.23 ± 0.03	0.90 ± 0.07	0.81 ± 0.05	0.75 ± 0.04	0.84 ± 0.04
Annual	-0.10 ± 0.04	-0.10 ± 0.05				

[a]Also shown are seasonal slopes (and associated slope uncertainty) for each individual site.

predicted versus measured CO will represent an upper limit provided there is not a systematic correlation between transport model bias and CO emission bias. Ideally, we would directly assess the bias of emission estimates and transport independently using precisely known emissions with passive tracers. However, the absence of such data on annual time scales over large regions precludes this, so we retain the assumption that systematic transport and emission error does not offset. This is partially supported by our results that demonstrated that improved simulation of surface wind speed, wind direction, and boundary height concurrently improved CO concentration estimates. Therefore, given the above approximation, we estimate the mean fractional bias and its uncertainty for regions and California wide as the weighted sum of slopes obtained across sites and seasons, where weights are estimated directly from the uncertainty in regression slope for each site and season shown in Table 3 using standard methods for propagation of errors [e.g., *Bevington*, 1969]. Combining results from CIT and SBC for SoCAB and WGC and ARV for the Central Valley, we estimate weighted annual mean slope biases of -0.10 ± 0.04 and -0.10 ± 0.05 for SoCAB and the Central Valley, respectively. Taking an upper bound on the uncertainty, these results indicate that on an annual basis the uncertainty associated with atmospheric inversions of CO across California is likely close to or smaller than 15%, as above, assuming the errors in the CO emissions maps are not correlated with errors in the transport model. This overall state-level uncertainty combining the results from individual sites assumes that the errors in the estimated best fit slopes are uncorrelated, and that the combined SFBay and CV emissions are approximately equivalent to those from SoCAB. This is supported by CARB 2012 CO annual emission estimates (ratio of SFBay + CV to SoCAB is 0.99). However, we note that there is possibility that the errors are correlated either expanding the overall uncertainty or reducing it due to anticorrelation.

Acknowledgments
We thank Dave Field, Dave Bush, Edward Wahl, Ken Reichl, Toby Walpert, and particularly Jon Kofler for assistance with measurements at WGC; John Lin, Christoph Gerbig, Steve Wofsy, Janusz Eluszkiewicz, and Thomas Nehrkorn for sharing the STILT code and advice; Paul Novelli for CO measurements used to estimate the CO background; Ying-Kuang Hsu, Bart Croes, Jorn Horner, Abhilash Vijayan, Vernon Hughes, and Webster Tassat for sharing CARB CO emission maps; and Krishna Muriki for assistance running the WRF-STILT models on the LBNL-Lawrencium cluster. Mixtli Campos-Pineda thanks a UC MEXUS-CONACYT Doctoral Fellowship. The measured and predicted CO data used in the analysis are shown in Figures 4–7 by season, and CO emission maps were obtained from the California Air Resources Board. This study was supported by the California Air Resources Board Research Division and the Natural Gas Research Program of the California Energy Commission under U.S. Department of Energy contract DE-AC02-05CH11231.

References

Andrews, A. E., et al. (2014), CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts, *Atmos. Meas. Tech.*, *7*(2), 647–687, doi:10.5194/amt-7-647-2014.

Angevine, W. M., L. Eddington, and K. Durkee (2012), Meteorological model evaluation for CalNex 2010, *Mon. Weather Rev.*, *140*, 3885–3906, doi:10.1175/MWR-D-12-00042.1.

Angevine, W. M., J. Brioude, and S. McKeen (2014), Uncertainty in Lagrangian pollutant transport simulations due to meteorological uncertainty from a mesoscale WRF ensemble, *Geosci. Model Dev.*, *7*, 2817–2829, doi:10.5194/gmd-7-2817-2014.

Bagley, J. E., and J. Miller (2015), Biophysical impacts of climate-smart agriculture in the Midwest United States, *Plant Cell Environ.*, doi:10.1111/pce.12485.

Bao, J.-W., S. A. Michelson, and P. Persson (2008), Observed and WRF-simulated low-level winds in a high-ozone episode during the Central California Ozone Study, *J. Appl. Meteorol. Climatol.*, *47*, 2372–2394, doi:10.1175/2008JAMC1822.1.

Bevington, P. R. (1969), *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Bianco, L., and J. M. Wilczak (2002), Convective boundary layer depth: Improved measurement by Doppler radar wind profiler using fuzzy logic methods, *J. Atmos. Oceanic Technol.*, *19*, 1745–1758, doi:10.1175/1520-0426(2002)019<1745:CBLDIM>2.0.CO;2.

Bianco, L., J. M. Wilczak, and A. B. White (2008), Convective boundary layer depth estimation from wind profilers: Statistical comparison between an automated algorithm and expert estimations, *J. Atmos. Oceanic Technol.*, *25*, 1397–1413, doi:10.1175/2008JTECHA981.1.

Bianco, L., I. V. Djalalova, C. W. King, and J. M. Wilczak (2011) Diurnal evolution and annual variability of boundary-layer height and its correlation to other meteorological variables in California's Central Valley, *Boundary Layer Meteorol.*, *140*, 491–511, doi:10.1007/s10546-011-9622-4.

Brioude, J., et al. (2013), Top-down estimate of surface flux in the Los Angeles Basin using a mesoscale inverse modeling technique: Assessing anthropogenic emissions of CO, NO_x and CO_2 and their impacts, *Atmos. Chem. Phys.*, *13*, 3661–3677, doi:10.5194/acp-13-3661-2013.

Dye, T. S., C. G. Lindsey, and J. A. Anderson (1995), Estimates of mixing depth from "boundary layer" radar profilers Preprints from the 9th Symposium on Meteorological Observations and Instrumentation, Charlotte, NC, March 27-31, 156-160 (STI-94212-1451).

Fiore, A. M., D. J. Jacob, I. Bey, R. M. Yantosca, B. D. Field, A. C. Fusco, and J. G. Wilkinson (2002), Background ozone over the United States in summer: Origin, trend, and contribution to pollution episodes, *J. Geophys. Res.*, *107*(D15), 4275, doi:10.1029/2001JD000982.

Gerbig, C., S. Koerner, and J. C. Lin (2008), Vertical mixing in atmospheric tracer transport models: Error characterization and propagation, *Atmos. Chem. Phys.*, *8*(3), 591–602.

Harding, K. J., and P. K. Snyder (2012), Modeling the atmospheric response to irrigation in the Great Plains. Part I: General impacts on precipitation and the energy budget, *J. Hydrometeorol.*, *13*, 1667–1686, doi:10.1175/JHM-D-11-098.1.

Holloway, T., C. Voigt, J. Morton, and S. N. Spak (2012), An assessment of atmospheric mercury in the Community Multiscale Air Quality (CMAQ) model at an urban site and a rural site in the Great Lakes Region of North America, *Atmos. Chem. Phys.*, *12*, 7117–7133, doi:10.5194/acp-12-7117-2012.

Jeong, S., C. Zhao, A. E. Andrews, L. Bianco, J. M. Wilczak, and M. L. Fischer (2012a), Seasonal variation of CH_4 emissions from central California, *J. Geophys. Res.*, *117*, D11306, doi:10.1029/2011JD016896.

Jeong, S., C. Zhao, and A. E. Andrews (2012b), Seasonal variations in N_2O emissions from central California, *Geophys. Res. Lett.*, *39*, L16805, doi:10.1029/2012GL052307.

Jeong, S., Y.-K. Hsu, A. E. Andrews, L. Bianco, P. Vaca, J. M. Wilczak, and M. L. Fischer (2013), A multitower measurement network estimate of California's methane emissions, *J. Geophys. Res. Atmos.*, *118*, 11,339–11,351, doi:10.1002/jgrd.50854.

Jeong, S., et al. (2016a), Estimating methane emissions in California's urban and rural regions using multitower observations, *J. Geophys. Res. Atmos.*, *121*, 13,031–13,049, doi:10.1002/2016JD025404.

Jeong, S., et al. (2016b), Estimating methane emissions from biological and fossil-fuel sources in the San Francisco Bay Area, *Geophys. Res. Lett.*, *44*, doi:10.1002/2016GL071794.

Kueppers, L. M., M. A. Snyder, and L. C. Sloan (2007), Irrigation cooling effect: Regional climate forcing by land-use change, *Geophys. Res. Lett.*, *34*, L03703, doi:10.1029/2006GL028679.

Lin, J. C., and C. Gerbig (2005), Accounting for the effect of transport errors on tracer inversions, *Geophys. Res. Lett.*, *32*, L01802, doi:10.1029/2004GL021127.

Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *J. Geophys. Res.*, *108*(D16), 4493, doi:10.1029/2002JD003161.

Mesinger, F., et al. (2006), North American regional reanalysis, *Bull. Am. Meteorol. Soc.*, *87*, 343–360, doi:10.1175/BAMS-87-3-343.

Michelson, S. A., and J.-W. Bao (2008), Sensitivity of low-level winds simulated by the WRF model in California's Central Valley to uncertainties in the large-scale forcing and soil initialization, *J. Appl. Meteorol. Climatol.*, *47*, 3131–3149, doi:10.1175/2008JAMC1782.1.

Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting–stochastic time-inverted Lagrangian transport (WRF–STILT) model, *Meteorol. Atmos. Phys.*, *107*(1-2), 51–64, doi:10.1007/s00703-010-0068-x.

Nehrkorn, T., J. Henderson, M. Leidner, M. Mountain, J. Elusziekwics, K. McKain, and S. Wofsky (2013), WRF simulations of the urban circulation in the Salt Lake City area for CO_2 modeling, *J. Appl. Meteorol. Climatol.*, *52*, 323–340, doi:10.1175/JAMC-D-12-061.1.

Newman, S., et al. (2013), Diurnal tracking of anthropogenic CO_2 emissions in the Los Angeles basin megacity during spring 2010, *Atmos. Chem. Phys.*, *13*, 4359–4372, doi:10.5194/acp-13-4359-2013.

Pfister, G., P. G. Hess, L. K. Emmons, J. F. Lamarque, C. Wiedinmyer, D. P. Edwards, G. Pétron, J. C. Gille, and G. W. Sachse (2005), Quantifying CO emissions from the 2004 Alaskan wildfires using MOPITT CO data, *Geophys. Res. Lett.*, *32*, L11809, doi:10.1029/2005GL022995.

Pillai, D., C. Gerbig, R. Ahmadov, C. Rödenbeck, R. Kretschmer, T. Koch, R. Thompson, B. Neininger, and J. V. Lavrié (2011), High-resolution simulations of atmospheric CO_2 over complex terrain–representing the Ochsenkopf mountain tall tower, *Atmos. Chem. Phys.*, *11*, 7445–7464, doi:10.5194/acp-11-7445-2011.

Rogers, R. E., A. Deng, D. Stauffer, Y. Jia, S. T. Soong, S. Tanrikulu, S. Beaver, and C. Tran (2013), Application of the weather research and forecasting model for air quality modeling in the San Francisco Bay area, *J. Appl. Meteorol. Climatol.*, *52*, 1953–1973, doi:10.1175/JAMC-D-12-0280.1.

Sacks, W. J., B. I. Cook, N. Buenning, S. Levis, and J. H. Helkowski (2009), Effects of global irrigation on the near-surface climate, *Clim. Dyn.*, *33*, 159–175, doi:10.1007/s00382-008-0445-z.

Skamarock, W. C., and J. B. Klemp (2008), A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, *J. Comput. Phys.*, *227*, 3464–3485, doi:10.1016/j.jcp.2007.01.037.

Smith, A., N. Lott, and R. Vose (2011), The Integrated Surface Database Recent Developments and Partnerships, *Bull. Am. Meteorol. Soc.*, *92*(6), 704–708, doi:10.1175/2011BAMS3015.1.

Sorooshian, S., J. Li, K. Hsu, and X. Gao (2011), How significant is the impact of irrigation on the local hydroclimate in California's Central Valley? Comparison of model results with ground and remote-sensing data, *J. Geophys. Res.*, *116*, D06102, doi:10.1029/2010JD014775.

Stohl, A., et al. (2012), Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: Determination of the source term, atmospheric dispersion, and deposition, *Atmos. Chem. Phys.*, *12*, 2313–2343, doi:10.5194/acp-12-2313-2012.

Strach, R. G., B. L. Weber, A. S. Frisch, C. G. Little, D. A. Merritt, K. P. Moran, and D. C. Welsh (1987), The precision and relative accuracy of profiler wind measurements, *J. Atmos. Oceanic Technol.*, *4*, 563–571.

van Vuuren, D. P., et al. (2009), Comparison of top-down and bottom-up estimates of sectoral and regional greenhouse gas emission reduction potentials, *Energy Policy*, *37*, 5125–5139.

Wyngaard, J. C., and M. A. LeMone (1980), Behavior of the refractive index structure parameter in the entraining convective boundary layer, *J. Atmos. Sci.*, *37*, 1573–1585, doi:10.1175/1520-0469(1980)037<1573:BOTRIS>2.0.CO;2.

Yienger, J. J., M. Galanter, T. A. Hollo-way, M. J. Phadnis, S. K. Guttikunda, G. R. Carmichael, W. J. Moxim, and H. levy II (2000), The episodic nature of air pollution transport from Asia to North America, *J. Geophys. Res.*, *105*, 26,931–26,945, doi:10.1029/2000JD900309.

Zhao, C., A. E. Andrews, L. Bianco, J. Eluszkiewicz, A. Hirsch, C. MacDonald, T. Nehrkorn, and M. L. Fischer (2009), Atmospheric inverse estimates of methane emissions from Central California, *J. Geophys. Res.*, *114*, D16302, doi:10.1029/2008JD011671.